Protein Chemotaxonomy of Genus *Datura*. IV.[1] Amino Acid Sequence of *Datura* Ferredoxins Depends Not on the Species but the Section of *Datura* Plants from Which It Comes

Yoshiki MINO

Department of Pharmacognosy, Osaka University of Pharmaceutical Sciences, 2–10–65 Kawai, Matsubara, Osaka 580, Japan. Received February 16, 1995; accepted March 24, 1995

The complete amino acid sequences of [2Fe-2S] ferredoxins from *Datura quercifolia* (section *Stramonium*) and *D. fastuosa* (section *Dutra*) have been determined by automated Edman degradation of the entire Cm-protein and of the peptides obtained by tryptic digestion and CNBr treatment. The *D. quercifolia* and *D. fastuosa* ferredoxins exhibited identical amino acid sequences to *D. stramonium* (section *Stramonium*) and *D. metal* (section *Dutra*) ferredoxins, respectively. This result suggests that the amino acid sequence of *Datura* ferredoxins depends on the section but not the species of *Datura* plants.

Key words *Datura quercifolia*; ferredoxin; amino acid sequence; protein chemotaxonomy; *Datura fastuosa*; Solanaceae

I have proposed the term 'protein chemotaxonomy' for molecular taxonomy based on the primary structures of common plant proteins, instead of secondary metabolites. To elucidate this concept, I planned a series of studies on the genus *Datura*, in which taxonomical problems still remain, using ferredoxin (Fd) as a common protein. I have already reported the primary structures of three different *Datura* Fds[1-3]: *Datura stramonium* Fd (*Ds*-Fd) which is isolated from the two varieties of *D. stramonium* (*Stramonium* section), *D. metel* Fd (*Dm*-Fd) from *D. metel* and *D. innoxia* (*Dutra* section), and *D. arborea* Fd from *D. arborea*, one of tree *Datura* (*Brugmansia* section). These results suggested that 1) *D. stramonium* var. *stramonium* and *D. stramonium* var. *tatula*, at first named by Linneous as two distinct species, *i.e.*, *D. stramonium* and *D. tatula*, should be considered two varieties of a single species, 2) *D. metel* is very close to *D. innoxia* in taxonomic position, 3) *Brugmansia* plants like *D. arborea* are reasonably considered to be a section of the genus *Datura* but not a separate genus from *Datura*. The present report describes the amino acid sequences of Fds from *D. quercifolia* (*Stramonium* section) and *D. fastuosa* (*Dutra* section), and the comparison of these two with the amino acid sequences of other *Datura* Fds reported previously.[1-3] In addition, the chromatographic properties of Cm-Fds and peptides obtained by their tryptic digestion are described, providing valuable information for the identification of Fds from the same genus or section.

Experimental

Apparatus The isolation of Fds and their related substances were carried out on a Water HPLC system (model 510 pumps, M680 gradient controller, and M441 absorbance detector) equipped with a stainless steel column, μ-Bondaspher C_{18}-100 Å (3.9 mm i.d. × 15 cm). The purification of Fds was performed on a Waters W650 protein system using a Mono Q column (HR5/5, Pharmacia).

Materials *Datura quercifolia* and *D. fastuosa* (Yaechosen asagao in Japanese, yellow and purple corolla types) were cultivated in an herb garden at this university. *D. quercifolia* has pale lavender corollas 4 to 7 cm long and 2 cm wide, smooth capsules, deeply pinnate leaves, and a purple stem. The leaves and stem are slightly pubescent. *D. fastuosa* has pale yellow (or purple) corollas (double or triple blossom type), glabrous stem and leaves, and very short spines or tubercles on the double capsules.

Isolation of Ferredoxins Frozen fresh *Datura* leaves (0.8 kg) were homogenized with 1.5 l of 0.02 M Tris–HCl buffer, pH 7.5. The homogenate was filtered through 2 layers of cloth mesh. The extract of *Datura* leaves was centrifuged and the supernatant was chromatographed on anion exchange resin (Q Sepharose FF) and Sephacryl S100HR. Repetition of the chromatography yielded pure Fds. The detailed protocols have been described in the previous report.[2] Homogeneity of the isolated Fd was checked by analytical polyacrylamide gel electrophoresis (PAGE) in a 10—15% gradient gel.

Sequence Determination The amino acid sequence of the Fds was determined using a gas-phase protein sequencer by automated Edman degradation of Cm-Fd, its tryptic peptides, and peptides resulting from cleavages at methionine residues. The C-terminal analysis was done with carboxypeptidase Y. Detailed procedures and other methods were described in the previous report.[2]

Results and Discussion

Properties The UV-Vis spectra of *D. quercifolia* (*Dq*) and *D. fastuosa* (*Df*) Fds exhibit absorption maxima at 275, 285 (sh), 330, 420, and 465 nm, which are similar to those for other *Datura* Fds.[4] As expected, the other physico-chemical properties of these two also corresponded to the assignment of [2Fe-2S] type Fd.[5]

Figure 1 shows the elution profile of the mixture of *Dq*- and *Df*-Fds on a Mono Q column. *Df*-Fd was eluted from the column at a slightly shorter retention time, suggesting a difference in amino acid sequence between *Dq*- and *Df*-Fds. After carboxymethylation of cysteine

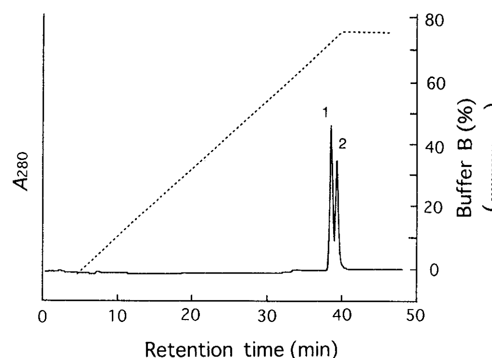

Fig. 1. HPLC of the Mixture of Fds from *Datura quercifolia* and *D. fastuosa*

1, *Datura fastuosa* Fd; 2, *D. quercifolia* Fd. Column, Mono Q (HR 5/5); flow rate, 0.5 ml/min; gradient elution of 0—0.75 M NaCl in 0.02 M Tris–HCl (pH 7.5) [buffer A, 0.02 M Tris–HCl (pH 7.5); buffer B, A + 1 M NaCl].

© 1995 Pharmaceutical Society of Japan

NII-Electronic Library Service

residues, individual S-carboxylmethylated (Cm)-Fds were not separated from each other on the reversed-phase HPLC using μ-Bondasphere C_{18} (data not shown). However, the peptide mixture obtained by the tryptic digestion of Cm-*Dq*-Fd exhibited a substantially distinct chromatogram from that for Cm-*Df*-Fd (Fig. 2). Of special interest is the fact that the peptide mixture which originated from *D. quercifolia* shows virtually the same profile as that for *D. stramonium* belonging to the same section "*Stramonium,*" and that the peptide mixture from *Df*-Fd corresponds exactly to that for *Dm*-Fd in the same section "*Dutra.*" This suggests two sets of identical primary structures between *Dq*- and *Ds*-Fds, and between *Df*- and *Dm*-Fds. However, the elution of peptides at the same retention time does not necessarily reveal an identical peptide. Thus, the amino acid composition of each peak fraction was analyzed. From the analytical results (Table 1), each tryptic peptide from *Dq*- or *Df*-Fd was clearly shown to be the same as the corresponding peptide from *Ds*- or *Dm*-Fd, respectively.

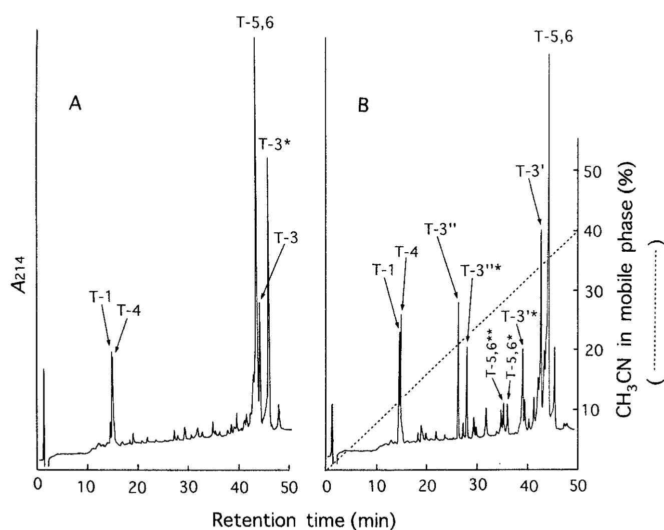

Fig. 2. HPLC of the Tryptic Digest of Each *Datura* Cm-Fds

A, *Datura fastuosa*; B, *D. quercifolia*. Column, μ-Bondasphere C_{18} (3.9 mm i.d. $\times$ 15 cm); flow rate, 1 ml/min; mobile phase, 0.1% TFA in H_2O/0.1% TFA in CH_3CN. T(1—6) represent peptides obtained from tryptic digestion. T-3′ and T-3″ are peptides obtained from only *D. fastuosa* Fd. The superscript * or ** reveals the peptides arising from the cleavage at the C-terminal side of Tyr or Phe residue, respectively.

All tryptic peptides, except for T-2, Val–Lys, were recovered. Note that the Lys-91 residue is trypsin-resistantive. In general, bigger peptides were easily retained on the hydrophobic stationary phase of the column, while T-4, 41—50, was eluted at relatively short retention times. This is because T-4 is a Cys- and Ser-rich peptide. It should also be noted that, in addition to Lys and Arg, cleavages were observed at some Tyr and Phe residues. For example, in the case of *Df*-Fd, four such peaks, T-3′* (7—23), T-3″* (28—37), T-5, 6* (51—73), and T-5, 6** (74—97), became bigger when the digestion time was extended with trypsin, while the original peaks, T-3′ (7—27), T-3″ (28—40), and T-5, 6 (51—97), became smaller. These chromatographic properties of tryptic peptides must be useful information for the identification of Cm-Fd from *Datura* plants.

Sequence Determination A summary of the sequence studies is shown in Fig. 3. The 42 steps of the automatic gas-phase Edman degradation method were successfully carried out using the Cm-*Dq*-Fd. Because of the tryptic-resistant Lys-91, the digestion of Cm-*Dq*-Fd with trypsin gave five peptides which were completely separated on a reversed-phase column (μ-Bondasphere C_{18}) in Fig. 2. Only T-2, Val–Lys, was missing. The sum of compositions of the 5 peptides agreed well with that of the original protein (Table 1). The T-4 (41—50), almost all parts of T-5, 6 (51—97), and CNBr-2 (70—97), were sequenced successfully. Carboxypeptidase Y released Gly (0.42), Thr (0.35), and Leu (0.20 mol/mol protein) subsequently from Cm-Fd after 15 min incubation, suggesting its C-terminal sequence to be Leu–Thr–Gly–COOH. The complete sequence composed of 97 residues was thus elucidated. As expected from the HPLC profile of the tryptic peptides, the primary structure of *Dq*-Fd was identical with that of *Ds*-Fd.

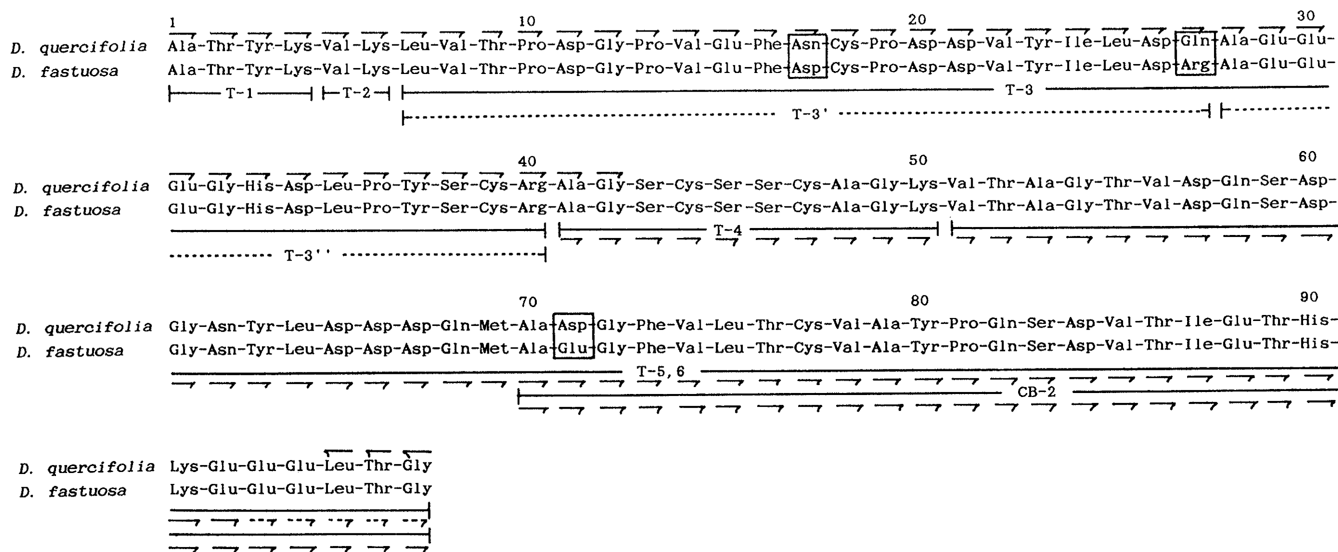

Fig. 3. Amino Acid Sequences of *Datura quercifolia* and *D. fastuosa* Fds

Arrows (⟶), (⟵) represent residues determined by automated Edman degradation and carboxypeptidase Y digestion, respectively. Dashed arrows (- - ⟶) indicate that the residue could not be unambiguously identified. T(1—6) represent peptides obtained from tryptic digestion. T-3′ and T-3″ are peptides obtained from only *D. fastuosa* Fd. CB-2 is a CNBr fragment. Three amino acid substitutions (boxed) were found in the amino acid sequence between *D. quercifolia* and *D. fastuosa* Fds.

NII-Electronic Library Service

Table 1. Amino Acid Composition[a] of the Peptides Isolated after Tryptic Digestion of *D. quercifolia* and *D. fastuosa* Cm-Fds

	D. quercifolia						*D. fastuosa*				
	T-1	T-3	T-4	T-5, 6	T-3*	T-4, 5, 6	T-3′	T-3″	T-5, 6	T-3′*	T-3″*
Residues covered	1—4	7—40	41—50	51—97	7—37	41—97	7—27	28—40	51—97	7—23	28—37
t_R	14.6	44.0	14.6	43.2	45.8	42.4	42.6	26.4	44.0	39.2	28.0
Cmc		1.8 (2)	1.9 (2)	1.0 (1)	0.6 (1)	2.2 (3)	0.9 (1)	0.8 (1)	1.0 (1)	0.5 (1)	
Asp (D)+Asn (N)		6.1 (6)		7.8 (8)	5.9 (6)	8.8 (8)	4.9 (5)	1.0 (1)	6.9 (7)	4.0 (4)	1.1 (1)
Thr (T)	1.0 (1)	1.1 (1)		5.1 (6)	1.1 (1)	5.2 (6)	1.1 (1)		5.4 (6)	1.0 (1)	
Ser (S)		1.3 (1)	2.7 (3)	2.2 (2)		5.2 (5)		1.0 (1)	2.1 (2)		
Glu (E)+Gln (Q)		5.2 (5)		6.7 (7)	5.2 (5)	7.2 (7)	1.2 (1)	2.8 (3)	7.7 (8)	1.1 (1)	2.9 (3)
Pro (P)		3.6 (4)		1.1 (1)	3.8 (4)		3.2 (3)	0.8 (1)	0.9 (1)	2.7 (3)	0.9 (1)
Gly (G)		2.5 (2)	2.2 (2)	4.3 (4)	2.7 (3)	7.0 (6)	1.1 (1)	1.3 (1)	3.9 (4)	1.2 (1)	1.1 (1)
Ala (A)	1.0 (1)	1.2 (1)	2.0 (2)	2.9 (3)	1.2 (1)	4.8 (5)		1.0 (1)	3.2 (3)		1.0 (1)
Val (V)		3.0 (3)		4.6 (5)	3.1 (3)	5.4 (5)	3.0 (3)		5.1 (5)	2.9 (3)	
Met (M)				0.6 (1)		0.6 (1)			0.7 (1)		
Ile (I)		1.0 (1)		0.9 (1)	1.0 (1)	1.4 (1)	0.9 (1)		1.1 (1)		
Leu (L)		3.0 (3)		2.8 (3)	3.1 (3)	3.6 (3)	2.0 (2)	1.0 (1)	3.1 (3)	1.0 (1)	1.0 (1)
Tyr (Y)	0.9 (1)	1.8 (2)		1.8 (2)	1.9 (2)	2.2 (2)	1.0 (1)	0.9 (1)	2.0 (2)	0.9 (1)	1.0 (1)
Phe (F)		1.0 (1)		1.0 (1)	1.9 (1)	1.2 (1)	0.9 (1)		1.0 (1)	1.0 (1)	
Lys (K)	1.0 (1)		1.0 (1)	0.9 (1)		2.0 (2)			1.9 (1)		
His (H)		1.0 (1)		0.9 (1)	1.9 (1)	1.2 (1)		1.0 (1)	1.0 (1)		0.9 (1)
Arg (R)		0.9 (1)					0.9 (1)	0.9 (1)			

a) Based on data obtained after 24 h hydrolysis. The numbers in parentheses are those calculated from each final sequence. * The peptides arising from the cleavage at the C-terminal side of Tyr residue.

Table 2. Matrix of Amino Acid Differences for *Datura* Fds

Section	Species	(1)	(2)	(3)
(1) *Stramonium*	*D. stramonium* var. *stramonium*	0	3	3
	D. stramonium var. *tatula*			
	D. quercifolia			
(2) *Dutra*	*D. metel*	3	0	4
	D. innoxia			
	D. fastuosa[a]			
(3) *Brugmansia*	*D. arborea*	3	4	0

a) Considered to be the synonymy of *D. metel*.

In the case of Cm-*Df*-Fd, because of additional Arg-27, two peptides, T-3′ (7—27) and T-3″ (28—40), were recovered from the tryptic digest instead of T-3 (7—40) for *Dq*-Fd. The sequencing experiments for *Df*-Fd, conducted in a similar manner, revealed a clearly identical amino acid sequence with that of *Dm*-Fd, with three amino acid differences from that of *Dq*-Fd. The Fds from both yellow- and purple-corolla types of *D. fastuosa* were also confirmed to have an identical amino acid sequence.

Taxonomic Consideration for *Datura* Species The genus *Datura* is divided into four sections: *Stramonium* as section I, *Dutra* as section II, *Ceratocaulis* as section III, and *Brugmansia* as section IV.[6] The species belonging to section I, *D. stramonium*, *D. ferox*, and *D. quercifolia* have erect flowers and capsules, which, when ripe, break regularly into four valves. The seeds are black. Six species belong to section II: *D. pruinosa*, *D. leichhardtii*, *D. meteloides*, *D. innoxia*, *D. discolor*, and *D. metel*. Flowers are erect, small or very large, five-lobed or ten-angled, white or colored, with a simple or double corolla. The capsules are nodding or inclined, breaking irregularly at maturity, and are spiny or tuberculate. The seeds are brown or black with a strophiole attached to the hilum. The *D. fastuosa* examined in this work is considered to be the synonymy of *D. metel*.[6] Table 2 shows the matrix of amino acid differences for the *Datura* Fds examined so far. As mentioned above, three amino acid differences were observed between *D. quercifolia* and *D. fastuosa*, belonging to different sections, while these Fds are identical with Fds from the same section, respectively. This result seems to be in good agreement with the taxonomic relationship.

Although few reports are available concerning the primary structures of Fds from plants that belong to the same genus, there are small differences among those of different species.[7,8] For example, two amino acid differences in their primary structures were detected in the case between *Phytolacca americana* and *P. esculenta* Fds. The Fds of *Equisetum telmateia* and *E. arvense* showed only one difference in primary structure. In the other hand, identical sequences have only been observed in the case of comparisons of varieties of the same species.[1,9] Recently, I have reported the identical sequence of Fd from different species of same section, *D. metel* and *D. innoxia* (both belong to the *Dutra* section) as a first example.[3] Also in the present work, *Dq*- and *Df*-Fds exhibited identical amino acid sequences for Fds from plants (*D. stramonium* and *D. metel*) that belong to the corresponding section, respectively. These results suggest that, at least in the case of genus *Datura*, the amino acid sequences of Fds depend not on the species but the section.

Acknowledgments The author is grateful to Prof. Yukio Noro, Faculty of Pharmacy, Meijo University, for supplying the seeds of *Datura quercifolia* and *D. fastuosa*. Thanks are also due to Emeritus Prof. Nagayo Ota, Dr. Seiji Inoue, and Prof. Kiyoshi Ikeda for their helpful discussion.

References

1) Part III: Mino Y., *Phytochemistry*, **37**, 429 (1994).
2) Mino Y., Usami H., Inoue S., Ikeda K., Ota N., *Phytochemistry*, **33**, 601 (1993).
3) Mino Y., *Phytochemistry*, **35**, 385 (1994).
4) Arnon D. I., Buchanan B. B., *Meth. Enzymol.*, **23**, 413 (1971).
5) Palmer G., "Iron-Sulfur Protein II," ed. by Lovenberg W.,

NII-Electronic Library Service

Academic Press Inc., New York, 1973, pp. 285—325.
6) Satina S., Avery A. G., "Blakeslee: Genus *Datura*," ed. by Avery A. G., Satina S., Rietsems J., Ronald Press, New York, 1959, pp. 16—47.
7) Matsubara H., Hase T., "Proteins and Nucleic Acid in Plant Systematics," ed. by Jansen U., Fairbrothers D. E., Springer, Berlin, 1983, pp. 168—181.
8) Huisman J. G., Stapel S., Muijsers A. O., *FEBS Lett.*, **85**, 198 (1978).
9) Shin M., Yokoyama Z., Abe A., Fukusawa H., *J. Biochem.*, **85**, 1075 (1979).

NII-Electronic Library Service